

# The ADM-Aeolus wind retrieval algorithms

By DAVID G. H. TAN<sup>1\*</sup>, ERIK ANDERSSON<sup>1</sup>, JOS DE KLOE<sup>2</sup>, GERT-JAN MARSEILLE<sup>2</sup>, AD STOFFELEN<sup>2</sup>, PAUL POLI<sup>3</sup>, MARIE-LAURE DENNEULIN<sup>3</sup>, ALAIN DABAS<sup>3</sup>, DORIT HUBER<sup>4</sup>, OLIVER REITEBUCH<sup>4</sup>, PIERRE FLAMANT<sup>5</sup>, OLIVIER LE RILLE<sup>6</sup> and HERBERT NETT<sup>6</sup>, <sup>1</sup>ECMWF, Shinfield Park, Reading RG2 9AX, UK; <sup>2</sup>KNMI, Postbus 201, 3730 AE, de Bilt, the Netherlands; <sup>3</sup>Météo-France, 42 avenue Coriolis, 31057 Toulouse, France; <sup>4</sup>DLR, Oberpfaffenhofen, D-82234 Wessling, Germany; <sup>5</sup>LMD/IPSL, 91128 Palaiseau, France; <sup>6</sup>ESA/ESTEC, Postbus 299, NL-2200-AG Noordwijk, the Netherlands

(Manuscript received 15 January 2007; in final form 20 August 2007)

## ABSTRACT

The ADM-Aeolus is primarily a research and demonstration mission flying the first Doppler wind lidar in space. Flexible data processing tools are being developed for use in the operational ground segment and by the meteorological community. We present the algorithms developed to retrieve accurate and representative wind profiles, suitable for assimilation in numerical weather prediction. The algorithms provide a flexible framework for classification and weighting of measurement-scale (1–10 km) data into aggregated, observation-scale (50 km) wind profiles for assimilation. The algorithms account for temperature and pressure effects in the molecular backscatter signal, and so the main remaining scientific challenge is to produce representative winds in inhomogeneous atmospheric conditions, such as strong wind shear, broken clouds, and aerosol layers. The Aeolus instrument provides separate measurements in Rayleigh and Mie channels, representing molecular (clear air) and particulate (aerosol and clouds) backscatter, respectively. The combining of information from the two channels offers possibilities to detect and flag difficult, inhomogeneous conditions. The functionality of a baseline version of the developed software has been demonstrated based on simulation of idealized cases.

## 1. Introduction

The Atmospheric Dynamics Mission, ADM-Aeolus, is the fourth of ESA's Earth Explorer Missions<sup>1</sup> (ESA, 1999; Stoffelen et al., 2005a). ADM-Aeolus is scheduled for launch in mid-2009 and has a projected lifetime of three years. Its objective is to demonstrate the capability to measure wind profiles from space using a Doppler Wind Lidar (DWL). The need for such data, with high accuracy and good vertical resolution, has been identified as a priority for the global observing system (WMO, 2004). The mission objectives and observation requirements have been designed to meet scientific goals in user communities in climate research, atmospheric modelling and numerical weather prediction (NWP). The polar orbit facilitates the global data coverage that is required, providing data also over the oceans which are currently poorly observed. The DWL will provide layer-averaged wind

measurements and observations<sup>2</sup> in 24 layers with configurable vertical distribution that can be modified in flight. The current baseline configuration will provide 1000 m vertical resolution through most of the atmosphere (from 2 to 16 km), 500 m below 2 km, and 2000 m between 16 and 26 km (Fig. 1).

The schematic in Fig. 1 shows the DWL instrument viewing from a low-altitude (~400 km) polar orbit in the direction perpendicular to the satellite track. Measurements are made in two receiver 'channels': Rayleigh for molecular returns and Mie for particulates. There is information on the horizontal line-of-sight (HLOS) wind component only (line-of-sight wind velocity divided by the cosine of the local elevation angle ~53°), which is close to east–west except at high latitudes. The unobserved wind component and the mass field will have to be statistically inferred within the data assimilation process (Riishøjgaard et al., 2004; Žagar, 2004; Stoffelen et al., 2005b; Žagar et al., 2005; Tan et al., 2007). The instrument will provide 50 km along-track average winds, separated by 150 km data gaps (Fig. 1); this is to ensure minimal error correlation between consecutive

\*Corresponding author.

e-mail: david.tan@ecmwf.int

DOI: 10.1111/j.1600-0870.2007.00285.x

<sup>1</sup>ESA Earth Explorers web site: [www.esa.int/esaLP/LPearthexp.html](http://www.esa.int/esaLP/LPearthexp.html).

<sup>2</sup>The term 'measurement' is used for instrument data characterized by horizontal scales of between 1 and 10 km, whereas 'observation' is used for aggregated data at horizontal scales of 50 km.

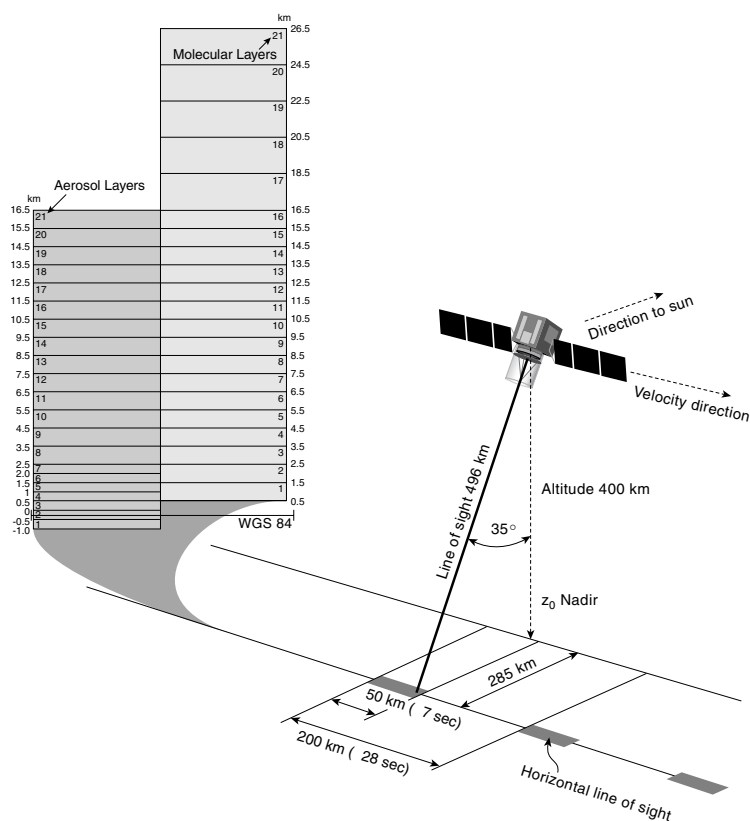


Fig. 1. Line-of-sight viewing geometry and proposed vertical distribution of the range bins (layers) for the ADM-Aeolus satellite, showing the aerosol (Mie) and molecular (Rayleigh) channels separately. Courtesy of ESA.

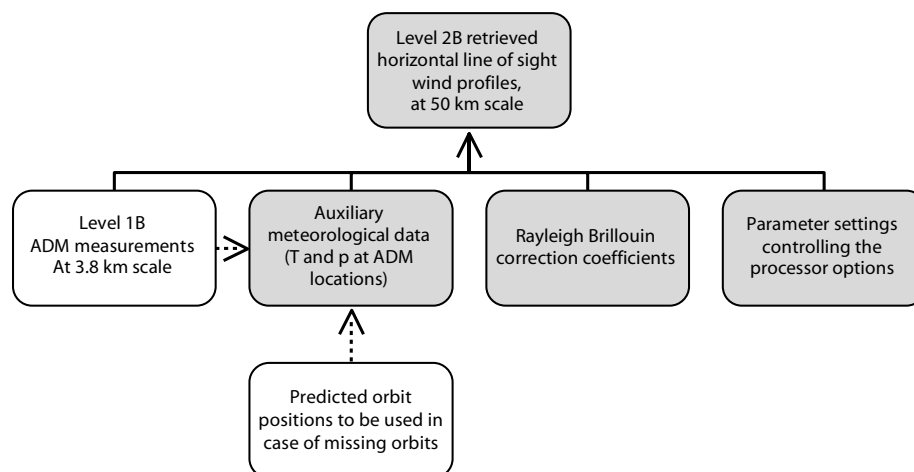


Fig. 2. Schematic showing main inputs to the ADM-Aeolus wind retrieval algorithm and the output L2B data. Unshaded boxes indicate that geolocation information is used to determine the locations of auxiliary meteorological data.

observations (Stoffelen et al., 2005a) and maximize the information content while conserving the energy consumption of the instrument. The accuracy of the ADM-Aeolus wind measurements and observations will depend primarily on the intensity of the backscattered laser light, which in the Mie channel depends on the presence and optical thickness of clouds, and the concentration of aerosol (Marseille and Stoffelen, 2003), and in the Rayleigh channel it depends mainly on the concentration of

molecules (i.e. the density of air) and attenuation by overlying aerosol and cloud. The expected yield and accuracy of Aeolus winds has been studied through detailed simulation (Tan and Andersson, 2005), based on model clouds (from the European Centre for Medium-Range Weather Forecasts, ECMWF) and climatological aerosol (Vaughan et al., 1995, 1998) distributions.

The literature cited above has noted that ADM-Aeolus offers substantial complementarity to existing wind observing

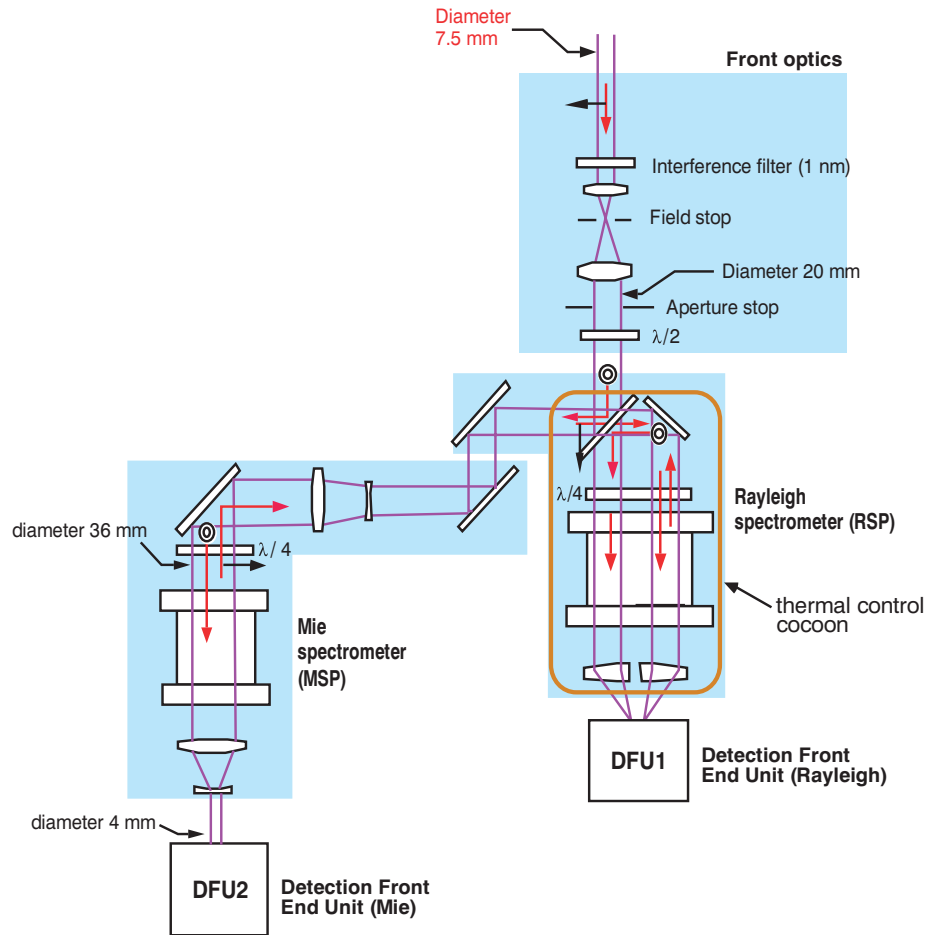


Fig. 3. Optical receiver architecture for the Aeolus DWL instrument. Courtesy of Astrium Satellites.

systems—to radiosondes, wind profilers and aircraft data by providing global coverage especially over oceans and away from the principal flight routes, and to atmospheric motion vectors by providing profiles with good vertical resolution. Complementarity to mass/temperature observing systems, that is, radiance and temperature data, has also been noted—this is regarded as particularly valuable for determining atmospheric motion on sub-synoptic scales and in the Tropics, that is, for regimes in which temperature data and conventional mass/wind balance relationships are inadequate (both empirically and from theoretical/dynamical arguments). Weissman and Cardinali (2007) showed that DWL observations taken in the North Atlantic from an airborne platform had a significant positive impact on analyses and forecasts of the ECMWF forecast system. Increasingly, simulated Aeolus data are being evaluated against real observations in NWP data assimilation/forecast experiments. For example, Tan et al. (2007) developed a technique based on the spread of an ensemble of data assimilations, to compare the expected impact of Aeolus data to that of the radiosonde and wind profiler network. They found that Aeolus can be expected to reduce analysis and short-range forecast uncertainty by an amount compa-

table to the radiosonde/wind profiler network, with the benefits being most apparent over oceans and in the Tropics. An underlying assumption of such studies is that the data processing chain, from raw instrument data up to Level-2B and including the generation of calibration/characterization data, is able to produce sufficiently accurate products (errors in HLOS wind estimates should be below  $2 \text{ ms}^{-1}$  throughout most of the atmosphere).

In this paper we describe the ADM-Aeolus Level-2B (L2B) wind retrieval algorithms which form part of the ADM-Aeolus data processing chain. The purpose of these algorithms is to obtain representative and accurate winds suitable for use in NWP. Level-1B (L1B) wind retrievals are not suitable for use in NWP for a number of reasons, the principal one being that L1B algorithms do not account explicitly for temperature and pressure effects on the response of the molecular (Rayleigh) channel of the instrument (see companion paper Dabas et al., 2008). The L2B algorithms use NWP information to take these effects into account. The design of the L2B algorithms takes account of the technical capabilities and constraints of the instrument, for example with respect to vertical and horizontal sampling, instrument pointing stability and zero wind calibration. Quality control and

product confidence indicators are important items that will be provided with the wind retrievals. In broken cloud scenes, it is envisaged that separate wind retrievals will be derived for clouds and clear air. This will be done through selective averaging of measurement-scale data in the layers of clear air above clouds, from cloud-top layers, from layers in and below thin clouds, and from layers with sufficient aerosol in the lower parts of the atmosphere.

The ADM-Aeolus is primarily a research and demonstration mission that will provide many opportunities for assessing the benefits of space-based wind profile information, and for defining the steps towards future operational DWL missions. Given the experimental nature of the mission, it has been recognized that data processing needs to have sufficient flexibility to explore the full potential of the mission data. The L2B wind retrieval algorithms discussed herein are likely to evolve during the mission. The evolution is expected to be relatively minor, but of course any changes will be thoroughly documented. The L2B software will be freely available to the meteorological community. The software has been designed to be portable, and specifically to run in three different contexts: (1) real-time processing at NWP centres with an interest to assimilate ADM-Aeolus winds within their own forecasting systems; (2) operational processing at the ECMWF to produce wind retrievals for delivery to ESA shortly after real time and (3) re-processing at ESA for situations in which delays in data delivery prevent processing within the ECMWF operational schedule, and to accommodate future algorithm improvements and upgrades.

The scope and purpose of the L2B wind retrieval processor is described in Section 2. Detailed descriptions of the algorithms are given in Section 3. Examples illustrating the behaviour of the retrieval, classification and error estimations are provided in Section 4, followed by concluding remarks in Section 5.

## 2. The scope and purpose of the Aeolus wind retrieval algorithms

The algorithms outlined in this paper are involved in calculating the L2B HLOS wind observations at the 50 km scale based on ADM-Aeolus measurements and instrument performance data. They were derived primarily to form part of a piece of software that creates the ADM-Aeolus Level-2B (L2B) data products. Based on the calibrated measurements (L1B) as inputs, they apply the modifications, corrections and additions required to obtain accurate and representative HLOS winds suitable for assimilation by NWP systems, as well as the appropriate quality control flags and uncertainty estimates. Key features of Aeolus products are summarized in Table 1. The so-called Level-2C (L2C) product is a superset of the L2B product and will be described elsewhere. Briefly, it contains additional output from ECMWF assimilation of L2B data, that is, ECMWF analysed winds at the Aeolus data locations. Thus, L2B products are intermediate between L1B and L2C data. Level-2A products (information

on aerosol and cloud layer optical properties) are described by Flamant et al. (2008).

The operational production of L2B data will be done at ECMWF slightly behind real time, just before the assimilation (and production of L2C) is carried out. The L2B processing uses a priori information on the state of the atmosphere at the time and place of the Aeolus L1B measurements. This information is best provided by the background fields of the NWP system, that is, the fields predicted by the forecast model run from the previous analysis. Meteorological background data, interpolated in the vertical plane along the flight track will also be created and delivered to ESA to facilitate re-processing of the Aeolus L1B data at a later time, and for off-line calibration tasks. Figure 2 is a schematic diagram showing the various data sets involved in creating the L2B data. It is envisaged that for their own purposes, many meteorological centres other than ECMWF will produce L2B data with local background inputs, and according to the timeliness constraints of their own operational NWP systems.

### 2.1. The ADM-Aeolus instrument

The payload of the ADM-Aeolus mission is a single DWL instrument. The instrument is a high-spectral resolution lidar operating in the ultraviolet, at wavelength  $\lambda_0 = 355$  nm. The Doppler frequency shifts  $\Delta\nu$  of the returned (elastic backscatter) atmospheric signals provide profile information on wind velocity along the instrument's line-of-sight  $v_{\text{LOS}}$ ,

$$\Delta\nu = -2 \frac{v_{\text{LOS}}}{\lambda_0} \quad (1)$$

while the signal amplitudes provide information on particle layers and their optical properties. The signal amplitudes also provide product confidence data, including error quantifiers, for the wind and particle information.

Although more details on the measurement principles underpinning the instrument are given in other papers from this volume (Dabas et al., 2008; Flamant et al., 2008) and elsewhere (Reitebuch et al., 2006), we summarize here the points most relevant for understanding the Level-2B wind retrieval algorithms. The solid curve in fig. 1 from Dabas et al. (2008) shows a nominal frequency spectrum measured by the instrument while the dashed curve shows how the spectrum is shifted in the presence of a  $50 \text{ ms}^{-1}$  LOS velocity. The spectral return from particles (aerosol, cloud) contributes the sharp narrow peaks and the spectral return from molecules contributes the broad portions (nearly Gaussian but modified by temperature and pressure effects). The instrument includes both a Mie receiver and a Rayleigh receiver designed to detect, respectively, the particulate and molecular return signals (Fig. 3). The Mie receiver is based on the fringe imaging technique with a Fizeau interferometer used in a mode where it forms an interference fringe whose spatial position is wavelength dependent, that is, a Doppler shift translates into a lateral displacement of the fringe. The Rayleigh receiver is based

Table 1. The main ADM-Aeolus data products

Product level	Description	Typical size (Megabytes per orbit)	Comments
Level 1B	Engineering-corrected HLOS winds	21–70	Near-real-time product. Spectrometer data at measurement scale, HLOS wind profiles using algorithms that do not account explicitly for scene classification nor for Rayleigh-Brillouin (pressure/temperature) effects.
Level 2A	Aerosol and cloud layer optical properties	7–10	Off-line product. See Flamant et al. (2008)
Level 2B	Meteorologically-representative HLOS winds	13–18	Shortly after near-real-time for operational products (generated at ECMWF), potentially near-real-time for other meteorological centres (depending on schedule). HLOS wind profiles using algorithms that (a) group measurements according to a scene-classification procedure and (b) account explicitly for Rayleigh-Brillouin effects—making use of NWP estimates of atmospheric temperature and pressure, typically from a short-range forecast. Subset of Level 2C products.
Level 2C	Aeolus-assisted wind vectors	19–24	Superset of Level 2B products. Adds ECMWF analysed winds (2 horizontal components) at the ADM-Aeolus locations, and supplementary product confidence data derived during assimilation of Level 2B data at ECMWF. The analysed winds take into account other atmospheric observations and the ECMWF forecast model through the data assimilation scheme.

on the double-edge technique with a sequential Fabry-Perot, itself including two band-pass filters ‘A’ and ‘B’ which produce two signal outputs that are then used in ecartometry mode to estimate the Doppler shifts. The two filters are centred on the edges of the backscattered molecular spectrum and placed symmetrically with respect to the laser central wavelength. Using the reflection on interferometers and polarization optics to perform the spectral separation, the atmospheric return signal first enters the Mie receiver and subsequently the Rayleigh receiver. The detector used for both channels is an Accumulation Charge Coupled Device (ACCD) with a 16 by 16 pixels useful image zone and optimized for ultraviolet sensitivity.

## 2.2. Measurements, observations and wind retrievals

L1B data are the instrumental input to the L2B processor. The L1B data set contains Aeolus measurements and observations and associated auxiliary parameters of one orbit, typically con-

sisting of about 200 independent observations. Both the L1B and L2B observation scale wind retrieval are computed from  $N$  Mie and Rayleigh measurements, each measurement itself being the result of the on-board analogue accumulation of  $P$  laser returns with an ACCD detector (see Fig. 4). In the standard operating mode of the Lidar,  $N = 14$  (up to 20 if laser warm-up pulses are considered) and  $P = 50$ , defining what is called the basic repeat cycle (BRC). Given the satellite ground-velocity of  $\sim 7.6 \text{ km s}^{-1}$ , and a pulse repetition frequency (PRF) of  $100 \text{ s}^{-1}$ , one measurement integrates the atmospheric return over a horizontal distance of  $(P = 50) \times (7.6 \text{ km s}^{-1}) / (\text{PRF} = 100 \text{ s}^{-1}) = 3.8 \text{ km}$ . The  $N$  contiguous measurements that make up one observation are representative of the wind field over a horizontal distance  $(N = 14) \times 3.8 \text{ km} \approx 50 \text{ km}$  (or  $\approx 70 \text{ km}$  if warm-up pulses are kept and  $N = 20$ ). Note that the starting points of two consecutive observations are separated by 200 km, that is to say there is a 150 km data gap between consecutive 50 km observations.

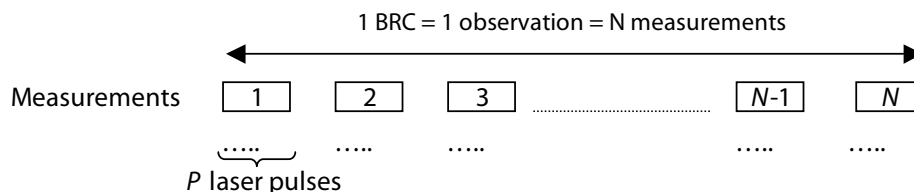


Fig. 4. Diagram showing the geometry of ADM-Aeolus measurements and observations. During one basic repeat cycle (BRC) of the laser,  $N$  measurements are acquired each resulting from the on-board analogue accumulation of  $P$  atmospheric returns generated by  $P$  laser pulses. One BRC corresponds to 50 km on the ground (Fig. 1).

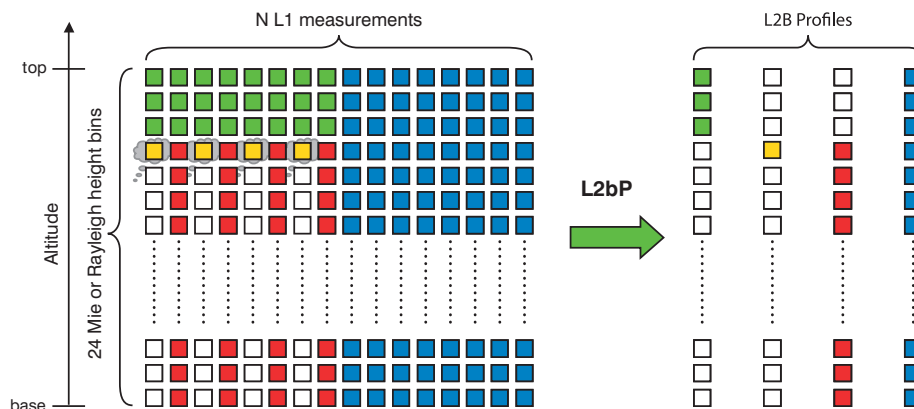


Fig. 5. Schematic illustration of the selective averaging of  $N$  measurements at 24 levels (or height bins, left) in a single Aeolus observation (over 50 km) into several partial or complete wind profile retrievals (right). The grey shading represents a broken cloud layer and the coloured squares represent classification of individual measurements in categories: green represents measurements above the broken cloud layer, yellow are the cloud top returns, red are between broken clouds, and blue are cloud free.

The L2B processor has access to the L1B data at measurement scale (3.8 km). This gives essential information on the heterogeneities of the atmosphere within one observation (50 km) and can help detect those situations which may lead to large measurement or representativeness errors in wind retrievals. In order to create representative averages within an observation, the measurements may be grouped into several categories each containing a profile of Mie and Rayleigh winds measured in similar homogeneous conditions.

An idealized meteorological situation that would result in several separate wind retrievals within a 50 km Aeolus L2B observation is illustrated in Fig. 5. The  $N$  Mie (or Rayleigh) measurements forming the observations are shown on the left. Single altitude, broken clouds can be seen in the left half, blocking some (in this case every second) of the measurements while others penetrate to the surface. The measurements in green above the cloud layer are averaged level-by-level to form the first retrieval containing only the three top-most levels. The cloud returns (yellow) are grouped into a second profile, in this case limited to a single level. The measurements from clear air between or below clouds (red) are averaged to form the third profile with valid winds from the surface to the altitude of the cloud layer. Finally, measurements from the second, homogeneous half of the scene (blue) produce the fourth wind profile retrieval.

### 2.3. L2B processing

In this section, we describe the effects and influences that need to be accounted for within the calculation of Aeolus L2B wind retrievals.

**2.3.1. Spatial coordinates.** The vertical coordinate of Aeolus L1B data is the height above the reference WGS84 ellipsoid whereas NWP models typically work with heights above the mean sea-level. The difference between the two altitudes may be several tens of meters, which is less than the thickness of range bins (250 m–2 km) but still cannot be neglected in contexts with strong vertical gradients in the wind. A conversion has thus been implemented which takes as reference the EGM96 geoid (see <http://cddisa.gsfc.nasa.gov/926/egm96/>).

**2.3.2. Cross-talk between the Mie and Rayleigh channels.** L2B processing of Mie channel data needs to account for contributions from the broad returned Rayleigh spectrum; such contributions are generally regarded as one component of background light. L2B processing of Rayleigh channel data needs to account for temperature and pressure dependence of the molecular backscatter (Rayleigh-Brillouin). It also needs to account for the possibility that reflections from the Mie channel's Fizeau interferometer enter the Rayleigh channel.

**2.3.3. Temperature and pressure effects.** L1B Rayleigh winds are obtained by converting the response of the dual Fabry-Perot

(defined by the difference of photons counted at the output of Fabry-Perot A and B divided by their sum) into a Doppler shift through the use of a response calibration curve. The true response curve depends on the atmospheric conditions in the sensing volume via the shape of the temperature- and pressure-dependent spectrum of the Rayleigh-Brillouin spectrum of the scattered light. In the presence of aerosol and clouds, Mie signal contaminates the Rayleigh signal, thereby modifying the response curve of the double Fabry-Perot. The Rayleigh winds must therefore be corrected for pressure, temperature and residual Mie light effects. Details of the correction scheme are presented by Dabas et al. (2008). As it requires prior knowledge on the actual temperature and pressure inside the sensing volume, it will necessarily be applied at L2B level which has access to NWP data. For each Rayleigh wind, the output will be

- (i) A Rayleigh HLOS wind corrected from pressure and temperature effects, using prior estimates  $P_{\text{ref}}$  and  $T_{\text{ref}}$ .
- (ii) The derivative  $a_R$  of the corrected wind with respect to the Rayleigh response (needed for error quantifier, see below).
- (iii) The derivative  $a_T$  of the corrected wind with respect to the temperature.
- (vi) The derivative  $a_p$  of the corrected wind with respect to the pressure.

These last two coefficients (sensitivities) permit optional further refinement of the correction when incrementally improved temperature and pressure estimates become available during NWP assimilation of atmospheric observations:

$$v_{\text{HLOS}}(T, P) = v_{\text{HLOS}}(T_{\text{ref}}, P_{\text{ref}}) + a_T(T - T_{\text{ref}}) + a_p(P - P_{\text{ref}}). \quad (2)$$

It is worth noting that the use of meteorological parameters (temperature and pressure estimates) in the L2B wind retrieval implies some correlation of errors between the wind estimates and the NWP model supplying the meteorological parameters. The correlation is directly proportional to the sensitivity coefficients reported in the L2B product, and is in fact rather small—as can be inferred from figs. 5 and 7 of the companion paper by Dabas et al. (2008).

The correction of the impact of Mie residue in the Rayleigh signals can be carried out at the same time. This requires information on the amount of Mie light at the receiver. This information cannot be obtained from the NWP model because the aerosol backscatter is not a model parameter and it cannot easily and reliably be related to existing model parameters: it must be estimated from the measured signals themselves, which in principle is possible by comparing the strength of the signals registered on the Mie and Rayleigh channel (the former one being mostly sensitive to particle backscatter while the latter one is mostly representative of the molecular return). Note that, due to the low sensitivity of the Rayleigh response to Mie contamination (see companion paper Dabas et al. 2008), the estimation of the rela-

tive contribution of particle backscatter does not need to be very accurate.

**2.3.4. Effect of atmospheric heterogeneities.** An essential task of the L2B processor is the identification of the atmospheric situations that are likely to produce large errors in the retrieved wind velocities. Cloud and aerosol structures are often spatially intermittent leading to potentially large signal differences from one measurement to the next at the same height. Measurements with relatively low signal-to-noise ratio can potentially deteriorate the wind estimate due to their high noise contributions. Here signal-to-noise ratio is defined as the useful signal divided by the detection noise associated with the useful signal and background light. Data points with strong Mie returns may be flagged because of the potential for the Mie returns to contaminate the detected signal in the Rayleigh channels. Large variability of strong Mie returns likely signals the presence of turbulence and vertical motion. In such cases, the optimal weighting of the strong and weak Mie and Rayleigh measurements has to be done with great care to ensure that spatially representative wind estimates are obtained.

Stratification of aerosol or cloud structures within a range bin (layer) may cause strong vertical gradients in the extinction profile resulting in vertically non-uniform contributions of laser light returning from within that range bin. In such conditions the retrieved wind would not represent the true mean wind over the layer. Without knowledge about the detailed stratification the wind retrieval can only be assigned to the mid-point of the layer, potentially incurring a significant height assignment error. However, it may be possible to flag the conditions in which this type of height assignment error is likely to occur, by careful scrutiny of differences in returns between subsequent range gates in the Mie and Rayleigh channels. Rayleigh signal height assignment problems could also be tackled by redistributing the available (24) Mie range bins to over-sample parts of the troposphere.

### 3. HLOS wind retrieval—the details

In this section we give details of the L2B wind retrievals from the Mie and Rayleigh channels, together with their associated error quantifiers. L2B products contain other ancillary parameters of a more technical nature, which will be described elsewhere. For the sake of brevity we do not include in this paper a number of preliminary steps related to the quality control and screening of Level-1B input data. Broadly speaking, these involve checking that the various parameters of interest are within reasonable bounds. Precise details, including the threshold values applied, can only be given after the mission's in-orbit commissioning phase.

In the introduction we referred to the need for flexibility in the data processing. This flexibility is manifested in the HLOS wind retrieval algorithms through their formulation in terms of weights  $w_{i,k}$  given to range gate (or layer)  $i$  of measurement  $k$ . The two indices are thus indicative of the two directions:  $i$  is in the vertical and  $k$  in the horizontal. For clarity of exposition, we present the

wind retrieval algorithms in terms of weights that are assumed given. We follow this with a baseline specification for how these weights are assigned for the examples presented in Section 4, and indicate what we consider to be the most promising options for alternative specifications to be explored during the mission lifetime.

### 3.1. L2B Mie channel HLOS wind estimate

The Mie-channel retrieved HLOS wind ( $v_{\text{HLOS}}$ ) is computed from the atmospheric return detected by the Mie channel ( $v_{\text{ATM}}$ ) and three correction terms. Two of these correction terms account for the laser internal reference path ( $v_{\text{REF}}$ ) and the satellite velocity relative to the ground ( $v_{\text{SAT}}$ ). The third correction is known as the ground wind ( $v_{\text{G}}$ ) correction term, and corrects for instrumental offsets that arise even for an atmospheric volume in which the wind velocity is zero, as expected for the range bins that intercept the Earth's surface. Consequently, the retrieved Mie HLOS wind is obtained by

$$v_{\text{HLOS}_i} = \frac{v_{\text{LOS}_i}}{\cos \varphi} \quad (3)$$

$$v_{\text{LOS}_i} = v_{\text{ATM}_i} - [v_{\text{REF}_i} + v_{\text{SAT}_i} + v_{\text{G}}], \quad (4)$$

where  $\varphi$  is the local elevation angle and  $i$  is the range-bin (or layer) index.

The atmospheric and internal reference LOS velocities,  $v_{\text{ATM}}$  and  $v_{\text{REF}}$ , are computed from the Doppler shift  $\delta$  by two similar expressions:

$$v_{\text{ATM}_i} = -\frac{1}{2}\lambda_0\delta_{\text{ATM}_i}/s_{\text{ATM}} \quad (5)$$

$$v_{\text{REF}_i} = -\frac{1}{2}\lambda_0\delta_{\text{REF}_i}/s_{\text{REF}}. \quad (6)$$

The factor  $-\lambda_0/2$ , with the wavelength  $\lambda_0 = 355$  nm, converts Doppler shifts into LOS wind velocities. The Doppler shift in frequency units is given by the ratio of the Doppler shift in CCD pixel units  $\delta$ , and an instrument response parameter  $s$  (to the first order equal to the frequency span  $\gamma = 1500$  MHz of the CCD, the so-called useful spectral range, divided by the useful number  $n = 16$  of CCD pixels). For the atmospheric path and the internal reference path these are denoted by subscripts ATM and REF, respectively. In an ideal instrument, the Doppler shift in pixel units would be the pixel position of the peak of the interference fringe output by the Fizeau interferometer and imaged onto the accumulation CCD, which we denote by  $\delta^m$ . However, to account for non-ideal non-linear response of the actual instrument, the Doppler shift in pixel units  $\delta$  is computed from  $\delta^m$  with a further non-linearity correction  $E$ :

$$\delta_{\text{ATM}_i} = \delta_{\text{ATM}_i}^m - E_{\text{ATM}}[\delta_{\text{ATM}_i}^m]. \quad (7)$$

$$\delta_{\text{REF}_i} = \delta_{\text{REF}_i}^m - E_{\text{REF}}[\delta_{\text{REF}_i}^m]. \quad (8)$$

Non-linearities are characterized regularly by a proper calibration procedure. It remains to specify how  $\delta_{\text{ATM}_i}^m$  and  $\delta_{\text{REF}_i}^m$  are computed. These are the result of applying the Mie core algorithm (Reitebuch et al., 2006) to the spectrometer readouts  $r$  (vector of the  $n = 16$  numbers of photons counted by the  $n = 16$  accumulation CCD pixels):

$$\delta_{\text{ATM}_i}^m = \text{Mie\_Core}[r_{\text{ATM}_i}] \quad (9)$$

$$\delta_{\text{REF}_i}^m = \text{Mie\_Core}[r_{\text{REF}_i}]. \quad (10)$$

The inputs to the Mie Core algorithm are the weighted spectrometer measurements (subscript  $k$ ) for the atmospheric return and the internal reference:

$$r_{\text{ATM}_i} = \sum_{k=1}^N w_{i,k} r_{\text{ATM}_{i,k}}; \quad r_{\text{REF}_i} = \sum_{k=1}^N w_{i,k} r_{\text{REF}_{i,k}}. \quad (11)$$

The spacecraft LOS velocity  $v_{\text{SAT}}$  is given by similarly weighting the measurement-scale velocities:

$$v_{\text{SAT}_i} = \sum_{k=1}^N w_{i,k} v_{\text{SAT}_{i,k}} \quad (12)$$

### 3.2. L2B Rayleigh channel HLOS wind retrieval

The L2B retrieved Rayleigh channel HLOS wind accounts for the atmospheric return detected by the Rayleigh channel and three correction terms. As in the case of Mie wind retrieval, these correction terms account for the laser internal reference path, the satellite velocity relative to the ground and the offsets that arise even for an atmospheric volume in which the wind velocity is zero. Consequently, the L2B Rayleigh HLOS retrieval is defined:

$$v_{\text{HLOS}_i} = v_{\text{ILIAD}_i} - \frac{1}{\cos \varphi} [v_{\text{REF}_i} + v_{\text{SAT}_i} + v_{\text{G}}]. \quad (13)$$

Here,  $v_{\text{ILIAD}}$  represents the HLOS wind retrieval taking into account the temperature and pressure dependence of the molecular backscatter (Rayleigh-Brillouin). The  $v_{\text{ILIAD}}$  velocity is obtained by applying the so-called ILIAD scheme to the Rayleigh channel measurements. ILIAD takes the general form:

$$v_{\text{ILIAD}_i} = \text{ILIAD}[R_{\text{ATM}_i}, \rho_i, T_i, p_i]. \quad (14)$$

Details of the ILIAD scheme are given in Dabas et al. (2008) so what is needed here is to specify the input parameters. These parameters denote weighted values within the scattering volume, for, respectively, the so-called instrument Rayleigh response  $R_{\text{ATM}_i}$ , the scattering ratio  $\rho = 1 + \beta_a/\beta_m$  (where  $\beta_a$  and  $\beta_m$  are the particle and molecular backscatter coefficients, respectively), the temperature  $T$  and the pressure  $p$ . The weighted values are obtained from a weighted average of the corresponding quantities at the measurement-scale:

$$\rho_i = \sum_{k=1}^N w_{i,k} \rho_{i,k}, \quad T_i = \sum_{k=1}^N w_{i,k} T_{i,k}, \quad p_i = \sum_{k=1}^N w_{i,k} p_{i,k}. \quad (15)$$



Similarly, the weighted Rayleigh response needed as input to ILIAD is given in terms of the weighted summation of useful signals in channels A and B (numbers of photons counted at the output of Fabry-Perot A and B corrected from background noise),  $N^A$  and  $N^B$ :

$$R_{\text{ATM}_i} = \frac{N_i^A - N_i^B}{N_i^A + N_i^B}, \quad (16)$$

where the weighted summation useful signal in channels A and B are:

$$N_i^A = \sum_{k=1}^N w_{i,k} N_{i,k}^A; \quad N_i^B = \sum_{k=1}^N w_{i,k} N_{i,k}^B. \quad (17)$$

Computation of a meteorologically weighted internal reference LOS velocity  $v_{\text{REF}}$  and the satellite's LOS velocity  $v_{\text{SAT}}$  are performed in an analogous manner, using the weights  $w_{i,k}$ .

### 3.3. Error quantifier for the L2B Mie channel HLOS wind estimate

The baseline for the L2B Mie error quantifier is the error standard deviation  $\sigma_{\text{MIE}}$  given by the equation:

$$\sigma_{\text{MIE}_i} = \frac{\lambda_0}{2 \cos \varphi} \left[ \sum_{j=3}^{18} \alpha_{i,j}^2 \right]^{-1} \sqrt{\sum_{j=3}^{18} \sigma_{i,j}^2 \alpha_{i,j}^2}, \quad (18)$$

where

$$\sigma_{i,j}^2 = \sum_{k=1}^N w_{i,k}^2 r_{\text{ATM}_{i,j,k}}^2 \quad (19)$$

is an estimate for the variance of the weighted sum of photo-counts  $r_{\text{ATM}}$  after removal of a detection chain (or analogue) offset, and

$$\alpha_{i,j} = \tau_j H \left[ \frac{1}{1 + \frac{4(f_j^+ - \hat{f}_i^{\text{MIE}})^2}{f_w^2}} - \frac{1}{1 + \frac{4(f_j^- - \hat{f}_i^{\text{MIE}})^2}{f_w^2}} \right]. \quad (20)$$

Here  $H$ ,  $f_w$  and  $\hat{f}_i^{\text{MIE}}$  (in frequency units) are the outputs of the Mie core algorithm ( $H$  is peak height and  $f_w$  is the full width at half maximum of the fitted Lorentzian spectrum, and  $\hat{f}_i^{\text{MIE}}$  is the Mie frequency estimate),  $\tau_j$  is a correction factor accounting for the obscuration of the primary mirror of the telescope by the tripod that bears the secondary mirror (characterized at ground before launch), and  $f_j^+$ ,  $f_j^-$  are the upper and lower frequencies of CCD pixel index  $j$ . With the useful spectral range  $\gamma = 1500$  MHz:

$$f_j^- = (j-3) \frac{\gamma}{16} - \frac{\gamma}{2} \quad \text{and} \quad f_j^+ = (j-2) \frac{\gamma}{16} - \frac{\gamma}{2}. \quad (21)$$

These expressions are derived as described in the Appendix. It is worth noting that the summation index in (18) runs from 3 to 18—indices 1 and 2 correspond to the CCD ‘pre-pixels’ which are not used for wind processing.

### 3.4. Error quantifier for the L2B Rayleigh channel HLOS wind estimate

The inversion step of the ILIAD scheme computes the horizontal line of sight wind  $v_{\text{HLOS}^\circ}$  using the sensitivity  $\partial v_{\text{HLOS}^\circ} / \partial p$ , for a zero scattering ratio,  $\rho = 0$ . For a given scattering ratio,  $\rho \neq 0$ , a linear correction is applied

$$v_{\text{HLOS}} = v_{\text{HLOS}^\circ} + \rho \frac{\partial v_{\text{HLOS}^\circ}}{\partial \rho}. \quad (22)$$

The radial wind  $v_{\text{HLOS}^\circ}$  is computed from the Rayleigh response  $R_{\text{ATM}}$  (see eq. 16), the temperature  $T$  and the pressure  $p$ . From this follows that the error on the calculated HLOS wind is

$$\Delta v_{\text{HLOS}} = \frac{\partial v_{\text{HLOS}^\circ}}{\partial R_{\text{ATM}}} \Delta R_{\text{ATM}} + \frac{\partial v_{\text{HLOS}^\circ}}{\partial T} \Delta T + \frac{\partial v_{\text{HLOS}^\circ}}{\partial p} \Delta p + \rho \frac{\partial v_{\text{HLOS}^\circ}}{\partial \rho}, \quad (23)$$

where  $\partial v_{\text{HLOS}^\circ} / \partial T$  and  $\partial v_{\text{HLOS}^\circ} / \partial p$  are the local sensitivities of the  $v_{\text{HLOS}^\circ}$  to  $T$  and  $p$ , respectively. For uncorrelated errors in scattering ratio, temperature, pressure and Rayleigh response, the resulting HLOS error,  $\sigma_{\text{HLOS}} = \sqrt{\langle \Delta v_{\text{HLOS}} \Delta v_{\text{HLOS}} \rangle}$  with  $\langle \rangle$  the operator for the expected covariance, is estimated as

$$\sigma_{\text{HLOS}} = \sqrt{\left( \frac{\partial v_{\text{HLOS}^\circ}}{\partial R_{\text{ATM}}} \sigma_{R_{\text{ATM}}} \right)^2 + \left( \frac{\partial v_{\text{HLOS}^\circ}}{\partial T} \sigma_T \right)^2 + \left( \frac{\partial v_{\text{HLOS}^\circ}}{\partial p} \sigma_p \right)^2 + \left( \frac{\partial v_{\text{HLOS}^\circ}}{\partial \rho} \sigma_\rho \right)^2}. \quad (24)$$

The sensitivities  $\partial v_{\text{HLOS}^\circ} / \partial \cdot$  are delivered as output from the ILIAD scheme. The dominant term stems from  $\partial v_{\text{HLOS}^\circ} / \partial R_{\text{ATM}}$ , which is about  $296 \text{ ms}^{-1}$ . The terms with the standard deviation of error for temperature and pressure, respectively  $\sigma_T$  and  $\sigma_p$  are small but this needs to be further tested. Equation (24) assumes that errors in the elevation angle  $\varphi$  are negligible, contributing a wind error of order  $0.01 \text{ ms}^{-1}$ .

It remains to provide an expression for the error standard deviation of the return signal  $R_{\text{ATM}}$ . From a sensitivity analysis it follows that the error in  $R_{\text{ATM}}$  is

$$\begin{aligned} \Delta R_{\text{ATM}} &= \frac{\partial R_{\text{ATM}}}{\partial N^A} \Delta N^A + \frac{\partial R_{\text{ATM}}}{\partial N^B} \Delta N^B \\ &= \frac{2N^B}{(N^A + N^B)^2} \Delta N^A - \frac{2N^A}{(N^A + N^B)^2} \Delta N^B. \end{aligned} \quad (25)$$

For independent errors in  $N^A$  and  $N^B$  the estimated error in the response  $R_{\text{ATM}}$ ,  $\sigma_{R_{\text{ATM}}} = \sqrt{\langle \Delta R_{\text{ATM}}^2 \rangle}$  is obtained by

$$\sigma_{R_{\text{ATM}}} = \frac{2}{(N^A + N^B)^2} \sqrt{N^{B^2} \sigma_A^2 + N^{A^2} \sigma_B^2}. \quad (26)$$

Recall from Section 3 that  $N^A$  and  $N^B$  are obtained from weighted sums of  $N_k^A$  and  $N_k^B$ , respectively, where each of the measurements in range gate  $i$  has a fractional weight  $w_{i,k}$ .  $\sigma_A$  and  $\sigma_B$ ,

at observation-scale, are thus as well obtained from a weighted sum, respectively

$$\sigma_A^2 = \sum_{k=1}^N w_{i,k}^2 \sigma_{A,k}^2 \quad \sigma_B^2 = \sum_{k=1}^N w_{i,k}^2 \sigma_{B,k}^2, \quad (27)$$

where  $\sigma_{A,k}$  and  $\sigma_{B,k}$  are the standard deviations of  $N_k^A$  and  $N_k^B$  (equal to the square root of  $N_k^A$  and  $N_k^B$  under the assumption that the photon count uncertainty is governed by Poisson statistics).

### 3.5. Selective averaging

Through aggregation of L1B data at measurement scale, L2B algorithms can produce multiple wind retrievals at observation-scale, as illustrated in Fig. 5. The procedure consists of two main steps: classification and weighting of each measurement-scale range gate.

(i) *Classification.* The scattering ratio estimate available in L1B products is an indicator of the presence of particles (cloud or aerosol) within the range-gate. The baseline classification exploits this by classifying all range-gates with scattering ratio above a threshold value as being ‘Cloudy’, and those below the threshold as ‘Clear’

(ii) *Weighting.* Let  $N_{\text{Cloudy}}$  denote the number of measurements for which range-gate  $i$  is classified as ‘Cloudy’, and let  $N_{\text{Clear}}$  denote the number of measurements for which range-gate  $i$  is classified as ‘Clear’. Then two sets of weights are defined:  $W_{\text{cloudy}}(i,k) = 1/N_{\text{Cloudy}}$  when range-gate  $i$  of measurement  $k$  is classified as ‘Cloudy’, and 0 otherwise.  $W_{\text{clear}}(i,k) = 1/N_{\text{Clear}}$  when range-gate  $i$  of measurement  $k$  is classified as ‘Clear’, and 0 otherwise. According to this simple weighting scheme, each measurement range-gate receives non-zero weight in just one set of weights.

Flexibility for future modifications includes classification based on L2B retrieval of optical parameters. The weights may be adjusted based on the signal-to-noise ratio of each measurement-scale range-gate.

## 4. Application to simulated data

The sensitivities of the baseline L2B algorithms to basic assumptions, errors and uncertainties in its main inputs have been tested through careful simulations and application to idealized cases. Two tests are described here to demonstrate retrieval of HLOS winds, scene classification with selective averaging, and estimation of wind retrieval uncertainty.

### 4.1. Scene classification and selective averaging

The purpose of this test is to assess the scene classification ability to identify clear from cloudy areas in an academic scenario and to group the measurement range-gates appropriately in or-

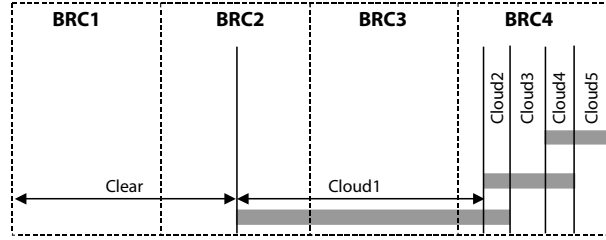


Fig. 6. Schematic representation of the test scenario with three cloud layers.

der to make L2B observations. The baseline scenario is made of four consecutive observations (or BRCs) where the HLOS wind is constant for all altitudes ( $50 \text{ ms}^{-1}$ ). The four BRCs contain a series of three cloud layers located at the altitudes 4, 9 and 15 km as illustrated in Fig. 6. The first BRC is clear (see Fig. 7 for the corresponding scattering ratio profile); the second is partly covered by Cloud1 featuring a single cloud layer at 4 km altitude; the third is entirely covered by Cloud1, and finally the fourth BRC contains five different cloud scenarios formed through a combination of partly overlapping clouds at three levels. The scenario called Cloud2 contains two cloud layers at 4 and 9 km altitudes, and so on. Briefly, the Cloud1, Cloud3 and Cloud5 BRCs contain one-layer clouds, while Cloud2 and Cloud4 contain two-layer clouds.

The scenario described above is run through the Aeolus End-to-End simulator (E2S v2.01) and L1B Processor (v1.05), developed in the frame of the satellite prime contract, in order to generate Level-0 and Level-1 instrument data. The Level-1 products are then used as input by the L2B processor (v1.2). The classification criterion uses the scattering ratio calculated by the L1B processor for each measurement range-gate with a threshold of 1.5 to separate clear from cloudy.

Upon running the L2B processor, the classification procedure assigns each measurement range-gate with three possible values:

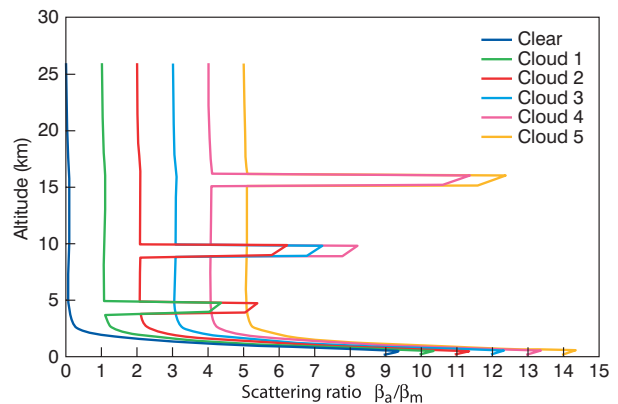
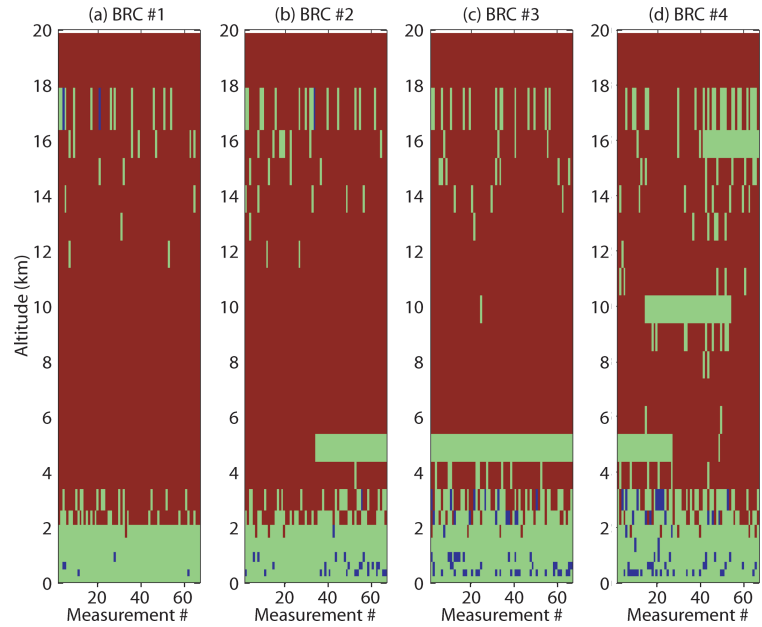


Fig. 7. Scattering ratio of the observations Clear, and Cloud1 to Cloud5. The cloud profiles have each been shifted by one unit in the  $x$ -direction, for convenience.

Fig. 8. Result of the L2B classification applied to the scenario shown in Fig. 6. Red (green, blue) indicates that the measurement range-gate is found clear (respectively: cloudy, not used).



clear, cloudy, or not used. The latter happens if for any reason the L1B input data are not within the expected range, such as below the surface or when measurement noise is too large. Figure 8 shows the result of the classification. Possibly due to noise, some

measurement range-gates are mistakenly identified as cloudy, but otherwise the cloud layer locations are found successfully.

We now evaluate the quality of the L1B and L2B wind retrievals in the various regions. Figure 9 shows the L1B HLOS

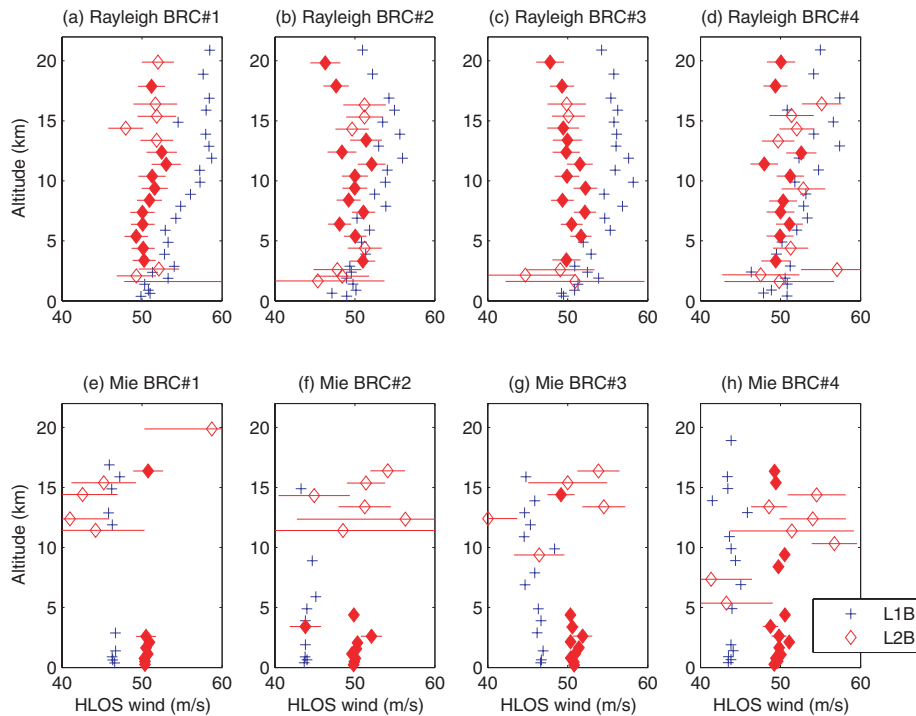


Fig. 9. Rayleigh (clear air, top) and Mie (from clouds, lower panels). L1B and L2B (see legend) HLOS wind retrievals at observation scale for the four BRCs of the scenario shown in Fig. 6. Horizontal bars extend up to  $\pm$  one L2B HLOS wind error standard deviation. Filled (empty) diamonds indicate where error estimates are smaller (larger) than  $2 \text{ ms}^{-1}$ .

observation wind retrievals (i.e. averaged over each BRC), obtained from both Rayleigh and Mie receivers. The Rayleigh wind retrievals show some apparent positive bias as compared to the true value of  $50 \text{ ms}^{-1}$  at higher altitudes, while the bias is negative for the Mie retrievals. The figure also shows the L2B Rayleigh retrievals for the clear observation and the L2B Mie retrievals for the cloudy observation. We retain only L2B data for which L2B quality flags are nominal, indicating no problem was encountered during L2B processing and L1B input data were found within expected range for either processing. Using proper meteorological temperature and pressure information, the bias in the Rayleigh retrievals is reduced by the L2B processor. Also, the error quantifiers provided by the L2B processor help identify the good-quality winds. For example, all L2B retrievals with small errors (i.e. error quantifiers  $< 2 \text{ ms}^{-1}$ , represented by filled diamonds on Fig. 9) are only found in clear regions for Rayleigh and in cloudy regions for Mie, as expected. Note that L2B retrievals for Rayleigh in cloudy regions are either absent because the L2B processor flagged these retrievals as poor or are suggested by the L2B to present large errors ( $> 2 \text{ ms}^{-1}$ , such as near the surface). Similarly, there are usually no good-quality L2B Mie retrievals in clear regions. Overall, the good L2B retrievals remain close to the true wind of  $50 \text{ ms}^{-1}$ .

The present test indicates that the L2B algorithms perform reasonably well with the academic scene presented here as regards classification and wind retrieval. The example shown here also illustrates that L2B products feature a combination of Mie and Rayleigh retrievals on which users should have more confidence thanks to the quality flags and representative error quantifiers. We anticipate that more investigation will be required to fine-tune the various parameters used for classification when actual data are collected with the ADM-Aeolus Airborne Demonstrator (Durand et al., 2006) and later with the spaceborne ADM-Aeolus itself during the spacecraft commissioning phase.

#### 4.2. Estimation of wind retrieval uncertainty

The purpose of this second test is to check the correctness of the proposed equation for estimating wind retrieval errors. The atmospheric scenario is simple: the vertical profiles of temperature and Rayleigh backscatter represent a mid-latitude winter case with altostratus clouds, but the wind is set artificially to zero, as is the aerosol and cloud backscatter and extinction (that is, no Mie return).

Based on this scenario, two BRCs were generated using the E2S. For the purpose of this test, all error-simulation options were switched off (satellite pointing and velocity errors...). The entire chain of ground segment processing tasks (L0→L1A→L1B) was applied. A total of 100 data files were generated with random Rayleigh CCD photon counts (Poisson's statistics). The L1B processor and the Rayleigh-Brillouin correction scheme were then applied to the 100 randomized data. The

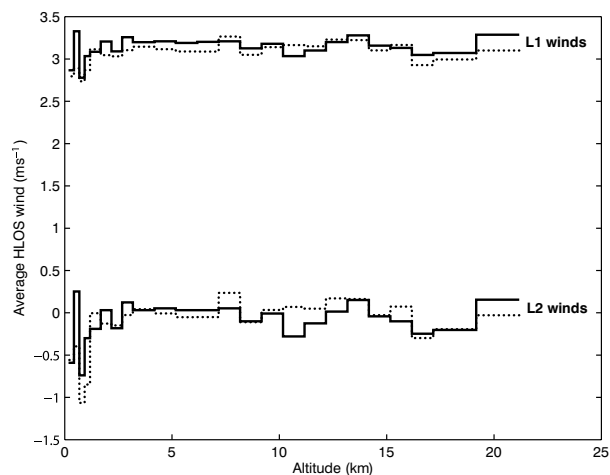


Fig. 10. Average of 200 L1B and L2B wind profiles computed from the randomized L1A data (solid curves, see legend). The true wind is zero. For L1 and L2 winds, the dotted curves show the wind profile obtained from the original L1A data file where no random Poisson statistics is applied.

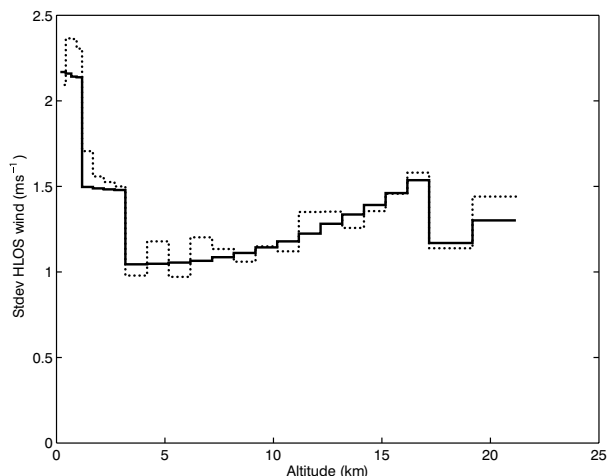


Fig. 11. Error standard deviation of random L2B winds (dotted) and errors (solid) estimated from the Rayleigh useful signals in A and B channels and sensitivity coefficients computed by the Rayleigh-Brillouin correction scheme.

average of the resulting 200 wind profiles (two BRCs per data file) is displayed in Fig. 10. It can be seen that L1 Rayleigh winds are biased while L2 winds are not. This confirms the necessity to correct Rayleigh winds from temperature effects and verifies the accuracy of the Rayleigh-Brillouin correction scheme. Figure 11 compares the actual standard deviation of L2 winds with the error standard deviation estimated by the L2B processor algorithm. In this simple scenario, the error estimation is given by eqs. (26) and (27). The results show very good agreement between estimated wind retrieval error, and the actual spread, which validates the error estimates.

## 5. Summary and conclusions

The ADM-Aeolus is primarily a research and demonstration mission flying the first Doppler wind lidar in space. Given the experimental nature of the mission, flexible data processing tools are being developed for use within ESA's operational ground segment and by the meteorological community. We have presented the algorithms developed for the processing of Aeolus data to Level-2B, that is, wind profile retrievals suitable for assimilation in NWP. The L2B processor provides a flexible framework for classification and weighting of measurement-scale (1–10 km) data into aggregated, observation-scale (50 km) wind profiles for assimilation. The main remaining scientific challenge is to produce representative winds in inhomogeneous atmospheric conditions, such as strong wind shear, broken clouds, and aerosol layers. The Aeolus instrument provides separate measurements in Rayleigh and Mie channels, representing molecular (clear air) and particulate (aerosol and clouds) backscatter, respectively. The combining of information in the two channels offers possibilities to detect and flag difficult, inhomogeneous conditions.

The functionality of a baseline version of the L2B processor has been demonstrated in terms of classification and wind retrieval. The corresponding computed error estimates of the retrieved winds have been validated. The next step is to apply the algorithms to real data obtained from an airborne Aeolus instrument demonstrator (Durand et al., 2006). Further refinement of the processor will continue even after launch of the satellite, in particular as based on results from the commissioning phase immediately after launch.

The L2B software is portable to a range of computers. It will be made freely available to the meteorological community. Operational Aeolus products will be available from ESA/ESRIN.

In Section 2.3.3, we noted that the use of meteorological parameters (temperature and pressure estimates) in the L2B Rayleigh channel wind retrieval implies some (small) sensitivity of the wind estimates to errors in the NWP model supplying the meteorological parameters. NWP models undergo regular improvements thereby changing (reducing) their error characteristics and to a lesser extent those of the L2B wind retrievals derived from them. For climate applications, where long-term trends are extracted from noisy signals, it is often desirable to remove such variations and so a re-processing capability is considered valuable. This could be achieved by including Level 2B processing of data from Aeolus, and any follow-on missions, within re-analysis projects (e.g. Uppala et al., 2005). It is easy to envisage other applications of Aeolus data, in particular to improve interpretation and use of other satellite data. This could include, but is not limited to, more accurate height assignment of atmospheric motion vectors and better detection of cloud-affected radiances. We welcome the participation of others to realize the full potential of the mission.

## 6. Acknowledgments

The authors thank the members of the ESA ADM mission advisory group for valuable comments and discussion. Comments from Anne-Grete Straume and two anonymous reviewers helped us to improve an earlier version of this paper. The work described here was funded by ESA contract 18555/04/NL/MM 'Implementation of Aeolus Level 2B/2C Processing Facility'.

## List of acronyms and terms

(A)CCD	(Accumulation) charge coupled device
ADM-Aeolus	Atmospheric Dynamics Mission (subsequently named 'Aeolus')
BRC	Basic repeat cycle
DWL	Doppler wind lidar
E2S	Aeolus End-to-End Simulator
EGM96	Earth geoid model, available from <a href="http://cddisa.gsfc.nasa.gov">http://cddisa.gsfc.nasa.gov</a>
ECMWF	European Centre for Medium-Range Weather Forecasts
ESA	European Space Agency
ESRIN	ESA Centre for Earth Observation
(H)LOS	(Horizontal) line of sight
ILIAD	Impact of Line shape on Aeolus Doppler estimates
L1A/B	Level-1A/B
L2A/B/C	Level-2A/B/C
NWP	Numerical weather prediction
PRF	Pulse repetition frequency
WGS84	World Geodetic System 1984, available from <a href="http://earth-info.nga.mil/GandG/wgs84/">http://earth-info.nga.mil/GandG/wgs84/</a>
WMO	World Meteorological Organization

## 8. Appendix: The Mie channel HLOS error estimate

The Mie core algorithm (Paffrath, 2006; Reitebuch et al., 2006) estimates the frequency of the Mie return by minimizing the cost function:

$$J(f) = \sum_{j=3}^{18} [N_{i,j} - \mu_j(f)]^2, \quad (\text{A1})$$

where  $N_{i,j}$  is a weighted photo-count for  $i$ th height-bin and  $j$ th CCD pixel and  $\mu_j(f)$  is a prediction of the same number based on an ad-hoc model (at present, a Lorentzian on top a uniform level of background light). The summation index in (A1) runs from 3 to 18 because indices 1 and 2 correspond to the CCD 'pre-pixels' which are not used for wind processing

The photo-counts  $N_{i,j}$  are computed from the measurement level CCD photo-counts  $r_{i,j,k}$  (after the removal of the detection chain offset estimated from CCD pixels 19 and 20 and correction

for tripod obscuration  $\tau_j$  as defined after eq. 20):

$$N_{i,j} = \sum_{k=1}^N w_{i,k} r_{i,j,k}, \quad (\text{A2})$$

where  $N$  is the total number of measurements and  $w_{i,k}$  are the weights allocated to measurement  $k$ .

In eq. (A1), the minimization bears only on the frequency, while four parameters are optimized by the Mie core algorithm (the central frequency, the amplitude, the width and the uniform level of background light). Here, we simplify the problem by limiting the optimization to the single frequency parameter, hoping that the equation we derive will apply with no major deficiency to the more complex case.

Let us denote by  $f_0$  the frequency that optimizes the cost function

$$\bar{J}(f) = \sum_{j=3}^{18} [\bar{N}_{i,j} - \mu_j(f)]^2, \quad (\text{A3})$$

where  $\bar{N}_{i,j}$  denotes the mathematical expectation of  $N_{i,j}$ . Then, let us approximate  $J(f)$  around  $f_0$  by a second order expansion:

$$J(f) \approx J(f_0) + \frac{\partial J}{\partial f}(f_0)(f - f_0) + \frac{1}{2} \frac{\partial^2 J}{\partial f^2}(f_0)(f - f_0)^2 \quad (\text{A4})$$

and assume that the frequency  $\hat{f}$  that minimizes  $J(f)$  is very close the frequency that minimizes (A3). This assumption can be written:

$$\hat{f} \approx f_0 - \left[ \frac{\partial^2 J}{\partial f^2}(f_0) \right]^{-1} \frac{\partial J}{\partial f}(f_0). \quad (\text{A5})$$

Now, let us denote  $J'(f) = J(f) - \bar{J}(f)$  and assume

$$\frac{\partial J'}{\partial f}(f_0) < 0 \text{ and } \frac{\partial^2 J'}{\partial f^2}(f_0) < 0. \quad (\text{A6})$$

Equation (A4) can then be approximated by the first order expansion

$$\begin{aligned} \hat{f} - f_0 &\approx - \left[ \frac{\partial^2 \bar{J}}{\partial f^2}(f_0) \right]^{-1} \left[ \frac{\partial \bar{J}}{\partial f}(f_0) + \frac{\partial J'}{\partial f}(f_0) \right] \\ &= - \left[ \frac{\partial^2 \bar{J}}{\partial f^2}(f_0) \right]^{-1} \frac{\partial J'}{\partial f}(f_0), \end{aligned} \quad (\text{A7})$$

where we have used  $\partial \bar{J} / \partial f(f_0) = 0$ . From (A7) follows that

$$\langle (\hat{f} - f_0)^2 \rangle = \left[ \frac{\partial^2 \bar{J}}{\partial f^2}(f_0) \right]^{-2} \left\langle \frac{\partial J'}{\partial f}(f_0) \frac{\partial J'}{\partial f}(f_0) \right\rangle. \quad (\text{A8})$$

From (A1) and (A3), we can write

$$J'(f) = \sum_{j=3}^{18} N'_{i,j} + 2 \sum_{j=3}^{18} N'_{i,j} [\bar{N}_{i,j} - \mu_j(f)] \quad (\text{A9})$$

so

$$\frac{\partial J'}{\partial f}(f_0) = 2 \sum_{j=3}^{18} N'_{i,j} \alpha_j(f_0), \quad (\text{A10})$$

where  $\alpha_j(f_0) = \partial \mu_j(f_0) / \partial f$ . Since  $\langle N'_{i,k} N'_{i,j} \rangle = \sigma_{i,j}^2 \delta(j - k)$  (the random fluctuations of the photo counts are independent), it follows that

$$\left\langle \frac{\partial J'}{\partial f}(f_0) \frac{\partial J'}{\partial f}(f_0) \right\rangle = 4 \sum_{j=3}^{18} \sigma_{i,j}^2 \alpha_j^2(f_0). \quad (\text{A11})$$

From eq. (A3)

$$\frac{\partial^2 \bar{J}}{\partial f^2}(f_0) = -2 \sum_{j=3}^{18} \frac{\partial \alpha_j}{\partial f}(f_0) [\bar{N}_{i,j} - \mu_j(f_0)] + 2 \sum_{j=3}^{18} \alpha_j^2(f_0). \quad (\text{A12})$$

The first term can be neglected if  $\bar{N}_{i,j} \approx \mu_j(f_0)$ . This condition should be met as long as the model  $\mu_j(f)$  is a good model for the photo-counts  $\bar{N}_{i,j}$ , so we can make the approximation

$$\frac{\partial^2 \bar{J}}{\partial f^2}(f_0) = 2 \sum_{j=3}^{18} \alpha_j^2(f_0). \quad (\text{A13})$$

Now, combining (A11) and (A13) gives

$$\langle (\hat{f} - f_0)^2 \rangle \approx \left[ \sum_{j=3}^{18} \alpha_j^2(f_0) \right]^{-2} \sum_{j=3}^{18} \sigma_{i,j}^2 \alpha_j^2(f_0). \quad (\text{A14})$$

This is the basis for the expression given in Section 3.1, which uses the photo-count model (integration over a CCD pixel of a Lorentzian plus a uniform level of background light)

$$\mu_j(f) = \tau_j \int_{f_j^-}^{f_j^+} \left\{ \frac{2A}{\pi f_w} \left[ 1 + \frac{4(x - f)^2}{f_w^2} \right]^{-1} + B \right\} dx, \quad (\text{A15})$$

where  $f_j^+$  and  $f_j^-$  are the upper and lower frequency bounds of CCD bin  $j$  as given by eq. (20) in the main paper,  $\tau_j$  is the tripod obscuration factor for bin  $j$ ,  $f_w$  is the full-width half maximum (FWHM) of the Lorentzian spectrum assumed for the Mie return,  $A$  is its amplitude, and  $B$  is the uniform level of background light. Considering this model, it follows that

$$\alpha_j(f_0) = \tau_j \frac{2A}{\pi \Delta} \left[ \frac{1}{1 + \frac{4(f_j^+ - f_0)^2}{f_w^2}} - \frac{1}{1 + \frac{4(f_j^- - f_0)^2}{f_w^2}} \right]. \quad (\text{A16})$$

In practice,  $A$ ,  $f_w$  and  $f_0$  are approximated with the parameters estimated by the Mie core algorithm (Reitebuch et al., 2006), the link between both sets being

$$\begin{aligned} \text{peak\_height} &\leftrightarrow \frac{2A}{\pi f_w} \\ \text{FWHM} &\leftrightarrow f_w \end{aligned} \quad (\text{A17})$$

frequency estimate  $\leftrightarrow f_0$

It now remains to give an expression for  $\sigma_{i,j}^2$ . If we assume that the random fluctuations of  $r_{i,j,k}$  follow independent, Poisson statistics, we have

$$\sigma_{i,j}^2 = \sum_{k=1}^{N_{\text{meas}}} w_{i,k}^2 \overline{r_{i,j,k}} \quad (\text{A18})$$

which we can approximate by

$$\sigma_{i,j}^2 \approx \sum_{k=1}^{N_{\text{meas}}} w_{i,k}^2 r_{i,j,k}^2.$$

## References

- Dabas, A., Denneulin, M.-L., Flamant, P., Loth, C., Garnier, A. and co-authors. 2008. Correcting winds measured with a Rayleigh Doppler lidar from pressure and temperature effects. *Tellus* **60A**, this issue.
- Durand, Y., Chinal, E., Endemann, M., Meynart, R., Reitebuch, O. and co-authors. 2006. ALADIN airborne demonstrator: a Doppler wind lidar to prepare ESA's Aeolus Explorer mission. *Proceedings SPIE Optics and Photonics* Volume 6296, 13–17 August 2006, San Diego.
- ESA, 1999. The four candidate Earth Explorer core missions—Atmospheric Dynamics Mission. *ESA Report for Mission Selection ESA SP-1233*(4), 145 pp.
- Flamant, P., Cuesta, J., Denneulin, M.-L., Dabas, A. and Huber, D. 2008. ADM-Aeolus retrieval algorithms for aerosol and cloud products. *Tellus* **60A**, this issue.
- Marseille, G.-J. and Stoffelen, A. 2003. Simulation of wind profiles from a space-borne Doppler wind lidar. *Q. J. R. Meteorol. Soc.* **129**, 3079–3098.
- Paffrath, U. 2006. *Performance Assessment of the Aeolus Doppler Wind Lidar Prototype*. PhD Thesis, Technical University Munich, DLR Forschungsbericht FB 2006–12, ISSN 1434–8454, 138 pp.
- Reitebuch, O., Paffrath, U., Huber, D. and Leike, I. 2006. Algorithm theoretical baseline document (ATBD): ADM-Aeolus level 1B products, AE-RP-DLR-L1B-001, issue 3.0, available from ESA.
- Riishøjgaard, L. P., Atlas, R. and Emmitt, G. D. 2004. The impact of Doppler lidar wind observations on a single-level meteorological analysis. *J. Appl. Meteor.* **43**, 810–820.
- Stoffelen, A., Pailleux, J., Källén, E., Vaughan, J. M., Isaksen, I. and co-authors. 2005a. The atmospheric dynamics mission for global wind measurement. *Bull. Amer. Meteorol. Soc.* **86**, 73–87.
- Stoffelen, A., Marseille, G.-J., Andersson, E. and Tan, D. G. H. 2005b. Comment on “The Impact of DWL Observations on a Single-Level Meteorological Analysis” by L. P. Riishøjgaard, R. Atlas, and G. D. Emmitt. *J. Appl. Meteorol.* **44**, 1276–1277.
- Tan, D. G. H. and Andersson, E. 2005. Simulation of the yield and accuracy of wind profile measurements from the Atmospheric Dynamics Mission (ADM-Aeolus). *Q. J. R. Meteorol. Soc.* **131**, 1737–1757.
- Tan, D. G. H., Andersson, E., Fisher, M. and Isaksen, L., 2007. Observing system impact assessment using a data assimilation ensemble technique: application to the ADM-Aeolus wind profiling mission. *Q. J. R. Meteorol. Soc.* **133**, 381–390.
- Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., da Costa Bechtold, V. and co-authors. 2005. The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.* **131**, 2961–3012.
- Vaughan, J. M., Brown, D. W., Nash, C., Alejandro, C. B. and Koenig, G. G. 1995. Atlantic atmospheric aerosol studies 2. Compendium of airborne backscatter measurements at 10.6  $\mu\text{m}$ . *J. Geophys. Res.* **100**(D1), 1043–1065.
- Vaughan, J. M., Geddes, N. J., Flamant, P. H. and Flesia, C. 1998. Establishment of a backscatter coefficient and atmospheric database, ESA contract reference 12510/97/NL/RE, Final Report. Available from ESA.
- Weissman, M. and Cardinali, C. 2007. Impact of airborne Doppler lidar observations on ECMWF forecasts. *Q. J. R. Meteorol. Soc.* **133**, 107–116.
- WMO 2004. *Proceedings of the 3<sup>rd</sup> WMO Workshop on “The Impact of Various Observing Systems on Numerical Weather Prediction”*, Alpbach, Austria, 9–12 March 2004 (eds H. Böttger, P. Menzel and J. Pailleux). *WMO/TD* No. 1228, 324 pp.
- Žagar, N. 2004. Assimilation of equatorial waves by line of sight wind observations. *J. Atmos. Sci.* **61**, 1877–1893.
- Žagar, N., Andersson, E. and Fisher, M. 2005. Balanced tropical data assimilation based on study of equatorial waves in ECMWF short-range forecast errors. *Q. J. R. Meteorol. Soc.* **131**, 987–1011.